

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 03-198099

(43)Date of publication of application : 29.08.1991

(51)Int Cl.

G10L 9/00

(21)Application number : 01-340106

(71)Applicant : NEC CORP

(22)Date of filing : 27.12.1989

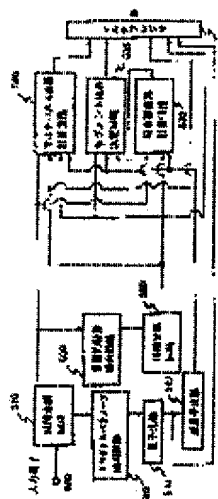
(72)Inventor : HANADA EISUKE

(54) VICE ENCODING SYSTEM

(57)Abstract:

PURPOSE: To express an audio signal satisfactorily with small amount of information by selecting the optimum sound source signal out of plural sound sources, and calculating the sound source signal by dividing the audio signal into segments of variable time length having continuous same feature.

CONSTITUTION: A time division circuit 510 divides an input audio signal into frames of time length decided in advance, and a spectrum parameter extraction circuit 520 finds a spectrum parameter, and a quantizer 530 quantizes the spectrum parameter, and a reverse quantizer 540 performs the reverse quantization of the parameter and outputs a result. An acoustic feature extraction circuit 550 extracts and outputs parameters representing a various kinds of acoustic characteristics from the input signal of the frame, and a feature classification circuit 560 performs decision by using the parameter. When a classified result shows a vowel-like signal, the feature classification circuit 560 stops the operation of a multi-pulse sound source selection circuit, and operates a vowel part sound source calculation circuit 600, and when the above result shows a non-vowel-like signal, the feature classification circuit 560 stops the operation of the vowel part sound source calculation circuit 600, and operates the multi-pulse sound source selection circuit 590. Thereby, it is possible to realize satisfactory sound with few quantity of computation.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

⑩ 日本国特許庁(JP)

⑪ 特許出願公開

⑫ 公開特許公報(A) 平3-198099

⑬ Int. Cl.⁵

G 10 L 9/00

識別記号

J

庁内整理番号

8622-5D

⑭ 公開 平成3年(1991)8月29日

審査請求 未請求 請求項の数 1 (全5頁)

⑮ 発明の名称 音声符号化方式

⑯ 特 願 平1-340106

⑰ 出 願 平1(1989)12月27日

⑱ 発 明 者 花 田 英 輔 東京都港区芝5丁目33番1号 日本電気株式会社内

⑲ 出 願 人 日本電気株式会社 東京都港区芝5丁目7番1号

⑳ 代 理 人 弁理士 内 原 晋

明 細 書

発明の名称 音声符号化方式

特許請求の範囲

離散的な音声信号を入力し、予め定められた時間長の区間(フレーム)に分割したのち分析したスペクトル包絡を表すスペクトルパラメータと音響的特徴を表すパラメータを抽出し、前記音響的特徴を利用して前記音声信号を分類して予め複数種類用意した音源モデルの中から最適なものを選択し、前記選択された音源モデルと前記スペクトルパラメータを用いて計算した現フレームの再生音声と前記入力音声との歪を計算し、現フレームの時間的に次のフレーム以後は前記の方法で計算された歪と1つ過去の区間について計算した歪との比を計算して、前記計算された比と予め定められた閾値とを比較することにより前記区間長に等しいかまたは長いセグメント長を求めて前記音声信号を分割し、前記選択された音源信号を表す符号と前

記計算されたスペクトルパラメータと前記セグメントの時間長と前記計算された音源信号を量子化して出力することとを特徴とする音声符号化方式。

発明の詳細な説明

(産業上の利用分野)

本発明は、音声信号を低いビットレートで高品質に符号化するための音声符号化方式に関する。

(従来の技術)

音声信号を低いビットレート、例えば16Kb/s程度以下で伝送する方式としては、マルチパルス符号化法などが知られている。これらは音源信号を複数のパルスの組合せ(マルチパルス)で表し、声道の特徴をデジタルフィルタで表し、音源パルスの情報とフィルタの係数を、一定時間区間(フレーム)毎に求めて伝送している。この方法の詳細については、例えばアラセキ、オザワ、オノ、オチアイ氏による“Multi-pulse Excited Speech Coder Based on Maximum Crosscorrelation Search Algorithm”, (GLOBECOM 83, IEEE Global Telecommunication, 講演番号23.3, 1983)(文献1)に記載されている。この

方法では、声道情報と音源信号を分離してそれぞれ表現すること、および音源信号を表現する手段として複数のパルス列の組合せ(マルチパルス)を用いることにより、複合後に良好な音声信号を出力することができる。

音声信号をより低いビットレートで伝送する方法としては、マルチパルス音源のピッチ毎の準周期性(ピッチ相関)を利用したピッチ予測マルチパルス法が提案されている。この方法の詳細は、例えば、特願昭58-139022号明細書(文献2)に詳しいのでここでは説明を省略する。

また、音声信号をさらに低いビットレートで伝送する方法としては、前記音源パルスの情報とフィルタの情報を一定時間毎ではなく、入力された音声信号を分析して得られる音響的特徴が持続している区間毎に区切り、(以下、音声信号を区切ったうえ分類する一連の処理をセグメンテーションといい、セグメンテーションの結果得られた各区間をセグメントと呼ぶ。)各セグメント毎に音源を求める方法をとることによって良好な再生

音声を得る方法をとることもできる。これらの方法については特願平1-23255号明細書(文献3)に詳しいのでここでは詳細を省略する。

(発明が解決しようとする課題)

しかしながら、文献1と文献2に掲載されている従来法ではビットレートが充分に高く音源パルスの数が充分なときは音質が良好であったが、ビットレートを下げて行くと音質が低下するという問題点があった。

また、前記文献3に掲載されている方法でも特に周期性を利用した音源を選択した場合でかつセグメントの長さが長くなった場合に音質が低下するという問題点があった。

本発明の目的は、ビットレートが高いところでも、下げていっても、また同じ音響的特徴を有する区間長が長くなっても、従来よりも良好な音声を少ない演算量で実現可能な音声符号化方式を提供することにある。

(課題を解決するための手段)

本発明による音声符号化方式は、離散的な音声信号を入力し、予め定められた時間中の区間(フレーム)に分割したのち分析スペクトル包絡を表すスペクトルパラメータと音響的特徴を表すパラメータを抽出し、前記音響的特徴を利用して前記音声信号を分類して予め複数種類用意した音源モデルの中から最適なものを選択し、前記選択された音源モデルと前記スペクトルパラメータを用いて計算した現フレームの再生音声と前記入力音声との歪を計算し、現フレームの時間的に次のフレーム以後は前記の方法で計算された歪と1つ過去の区間について計算した歪との比を計算して、前記計算された比と予め定められた閾値とを比較することにより前記区間長に等しいかまたは長いセグメント長を求めて前記音声信号を分割し、前記選択された音源信号を表す符号と前記計算されたスペクトルパラメータと前記セグメントの時間長と前記計算された音源信号を量子化して出力することの特徴とする。

(作用)

本発明による符号化方式において選択された音源モデルが適しているセグメントにおけるセグメントの区間長(以下、セグメントの区間長を略してセグメント長という。)の決定方法について第2図及び式を用いて説明する。第2図は、セグメント内の音響的特徴が母音性を示す場合である。

第2図は母音性を示すセグメントにおける音源信号計算回路及びセグメント境界決定回路のブロック図を示す。信号入力端子310からは予め定められた区間長(例えば10msec)(以下、第1フレームという)の信号 $x_1(n)$ が入力される。また、スペクトルパラメータ入力端子300からは前記第1フレームの入力信号 $x_1(n)$ から計算されたスペクトルパラメータが入力される。音響パラメータ入力端子305からは前記第1フレームの入力信号 $x_1(n)$ の音響的特徴を表すパラメータが入力される。音源選択回路306は前記入力された音響的特徴を表すパラメータを用いて予め複数個用意された音源モデルから最適なモデルを選択する。具体的な方法は前記文献3を参照できる。前記音響的特徴が母音を示すときは、音

源計算回路320は前記第1フレームの入力信号 $x_1(n)$ と前記入力されたスペクトルパラメータを用いて前記選択された音源モデルを用いて音源信号を計算する。前記第1フレームに対する処理が終了すると、現在のセグメント長を前記区間長に設定する。バッファ330は前記計算された音源信号 $x_1(n)$ と現在のセグメント長を蓄積し、前記第1フレームの入力信号 $x_1(n)$ をそのまま出力する。また、再生フィルタ324は前記音源計算回路の出力である音源信号と前記入力された音響的特徴を表すパラメータを用いて音声信号 $s_1(n)$ を再生する。再生の具体的な方法は前記文献3を参照出来る。歪計算回路323は前記再生された音声信号 $s_1(n)$ と前記入力された第1フレームの音声信号 $x_1(n)$ を用いて前記第1フレームにおける歪 E_1 を予め定めた次式(1)のように計算する。

$$E_1 = d \{s_1(n), x_1(n)\} \quad (1)$$

前記歪計算回路の出力 E_1 は歪保存回路385に初期値として保存する。次に信号入力端子340から前記バッファ330の出力信号の時間的に次のフレームの

ファ330に蓄積し、前記歪計算回路370において計算された歪を前記歪保存回路に初期値として保存する。前記バッファは前記蓄積された音声信号を前記信号形成回路へ出力する。そして次のフレームの入力信号に対する処理を行う。一方、前記歪か前記閾値よりも大きい場合は、対象としている入力信号の開始点に境界があると判定し、バッファ330に蓄積されている音源信号とセグメント長を出力端子390に出力する。

以上示したような方法を用いることによって、前記文献2における方法に対しても大幅に計算量を削減した上、歪を良好に小さく保つことが可能である。

(実施例)

本発明の一実施例を示す第1図において、入力端子500から離散的な音声信号を入力する。時間分割回路510では前記入力された音声信号を予め定められた時間長(例えば10msec)のフレームに分割する。スペクトルパラメータ抽出回路520では前記フレームの音声信号のスペクトルを表すスペクトル

入力信号 $x_2(n)$ を入力する。信号形成回路345は前記バッファから出力された信号に続けて前記入力端子340からの入力信号を加えた区間長の入力信号 $x_2'(n)$ を作成して出力する。音源計算回路350は前記信号形成回路からの出力信号に対して前記選択された音源モデルを用いて音源信号を計算する。再生フィルタ360は前記入力されたスペクトルパラメータと前記計算された音源信号を用いて音声信号 $s_2'(n)$ を再生する。歪計算回路370は前記再生フィルタによって再生された再生信号 $s_2'(n)$ と入力信号 $x_2'(n)$ との歪 E_2 を前記式(1)と同じ計算方法で計算する。境界判定回路380は前記計算された歪 E_2 と前記歪保存回路に保存されている1つ前のフレームにおける歪 E_1 との比 R_{12} を次式(2)に従って計算し、得られた結果 R_{12} と予め定められた閾値 T とを比較する。

$$R_{12} = E_2 / E_1 \quad (2)$$

前記歪の比 R_{12} が前記閾値 T よりも小さい場合はセグメント長を更新し前記音源計算回路350の出力である音源信号と前記更新されたフレーム長をバッ

パラメータを、衆知のLPC分析法によって求める。量子化器530は求められたスペクトルパラメータを量子化する。逆量子化器540は、量子化されたスペクトルパラメータを逆量子化して出力する。

音響的特徴抽出回路550は、前記フレームの入力信号から種々の音響的特徴を表すパラメータを抽出して出力する。特徴分類回路560は、前記出力された音響的特徴を表すパラメータを用いて、前記フレーム内の入力信号が母音性の特徴を持っているかどうか判定する。判定に用いる音響的特徴としては例えばフレーム内のパワーまたはRMS、ピッチゲインなど衆知の方法によって求めることができるパラメータがある。

前記分類結果が母音性信号である場合は、前記特徴分類回路は前記マルチパルス音源選択回路の作動を停止し母音部音源計算回路600を作動させる。前記母音部音源計算回路は例えば小澤氏による“種々の音源を用いる4.8kb/s音声符号化方式(SPM EX)”(電子情報通信学会音声研究会資料SP-89-2 1989年、文献4)の中で用いられている改良ピッチ補

間マルチパルス音源を音源モデルとして用いて前記求めたセグメントの音源信号を計算して量子化して出力する。母音部の音源を計算し、セグメント境界決定回路605に出力する。前記セグメント境界決定回路は前記母音部音源計算回路の出力である音源信号と前記逆量子化器の出力と前記入力信号を用いて作用の項に示したような方法を用いてセグメント境界があるかどうか決定し、境界があると判定されれば前記母音音源計算回路において計算された音源信号を表す符号とセグメント長を出力する。

一方、前記分類結果が母音性信号でない場合は、前記特徴分類回路は母音部音源計算回路600の作動を停止しマルチパルス音源選択回路590を動作させる。前記マルチパルス音源計算回路590は、前記特徴分類回路が前記入力信号が母音性信号でないという分類結果を出力した場合に、例えば前記文献1に見られるような衆知の方法で音源パルスを計算する。

また、分類に応じて用いる音源は、上の例に示した以外に例えば前記文献4の中で用いられているように予め複数種類の音源モデルを用意して入力信号の音響的特徴によって最適な音源モデルを切り替えて用いる形とすることもできる。例えば母音性以外の信号の音源としては破裂、過渡性のときにマルチパルス音源、摩擦性のときに乱数コードブック音源を用いることができる。前者については前記文献1に、後者については例えばシェレーダー、アタル両氏による“CODE-EXCITED LINEAR PREDICTION (CELP): HIGH-QUALITY SPEECH AT VERY LOW BIT RATES” (ICASSP '85 講演番号25.1.1 1985)(文献5)と題した論文等に詳しいのでここでは説明を略す。

また、本発明によるセグメント長を決定するための方法は、上の例では母音性を示す場合のみに用いたが、これ以外の場合、例えば摩擦性などに用いてもよい。
(発明の効果)

前記量子化器の出力と前記特徴分類回路の出力と各音源計算回路の出力である音源符号を表す符号とセグメント中を表す符号はマルチプレクサ610の入力となる。前記マルチプレクサはこれらの入力を効率的に多重化して出力する。

以上述べた構成は本発明の一構成に過ぎず、種々の変形も可能である。

マルチパルスの計算方法としては、前記文献1に示した方法の他に、種々の衆知な方法を用いることができる。

また、スペクトルパラメータとしては、他の衆知なパラメータ(縦スペクトル対、ケプストラム、メルケプストラム、対数断面積比等)を用いることもできる。

また、セグメント長を決定するための閾値はどの場合でも1つの数値としてもよいし、セグメント長に応じて算出した数値としてもよい。

また、前記求める歪は例えば2乗誤差としてもよいし、この他衆知の距離尺度を用いてもよい。

本発明によれば、音源信号を予め用意した複数個の音源の中から最適なものを選択し、音声信号を同一の特徴の連続している可変時間長のセグメントに分割して前記音源信号を計算することにより、従来法に比べ少ない伝送情報量で音声信号を良好に表すことができるという大きな効果がある。

図面の簡単な説明

第1図は本発明による音声符号化方法の一実施例の構成を示すブロック図、第2図は本発明の作用を説明するための図である。

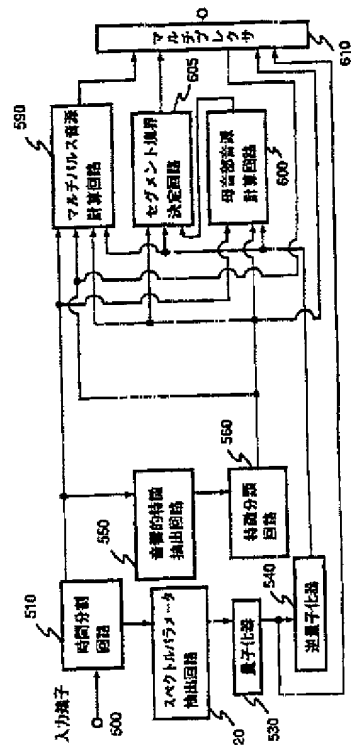
図において、

306…音源選択回路、320、350…音源計算回路、330…バッファ、380…境界判定回路、345…信号形成回路、324、360…再生フィルタ、323、370…歪計算回路、385…歪保存回路、510…時間分割回路、520…スペクトルパラメータ計算回路、530…量子化器、540…逆量子化器、550…音響的特徴抽出回路、560…特徴分類回路、590…マルチパルス計算回路、

600 母音部音源計算回路、605 セグメント境界決定回路、610 マルチプレクサ。

代理人 弁理士 内原 晋

第 1 図



第 2 図

